Vesa Närhi

The Use of Clinical Neuropsychological Data in Learning Disability Research

JYVÄSKYLÄ STUDIES IN EDUCATION, PSYCHOLOGY AND SOCIAL RESEARCH 193

Vesa Närhi

The Use of Clinical Neuropsychological Data in Learning Disability Research

Esitetään Jyväskylän yliopiston yhteiskuntatieteellisen tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston vanhassa juhlasalissa (S212)
maaliskuun 16. päivänä 2002 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Social Sciences of the University of Jyväskylä,
in Auditorium S212, on March, 2002 at o'clock noon.

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ 2002

The Use of Clinical Neuropsychological Data in Learning Disability Research

JYVÄSKYLÄ STUDIES IN EDUCATION, PSYCHOLOGY AND SOCIAL RESEARCH 193

Vesa Närhi

The Use of Clinical Neuropsychological Data in Learning Disability Research

JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ 2002

Editors
Tapani Korhonen
Department of Psychology, University of Jyväskylä
Pekka Olsbo and Marja-Leena Tynkkynen
Publishing Unit, University Library of Jyväskylä

Cover picture: Fiina Närhi

URN:ISBN:978-951-39-8071-9
ISBN 978-951-39-8071-9 (PDF)
ISSN 0075-4625

ISBN 951-39-1119-5
ISSN 0075-4625

Jyväskylä University Printing House, Jyväskylä
and ER-Paino Ky, Lievestuore 2002

ABSTRACT

Närhi, Vesa
The use of clinical neuropsychological data in learning disability research
Jyväskylä: University of Jyväskylä, 2002, 103 p.
(Jyväskylä Studies in Education, Psychology and Social Research,
ISSN 0075-4625; 193)
ISBN 951-39-1119-5
Yhteenveto: Asiakastyön yhteydessä kerätyn neuropsykologisen aineiston käyttö oppimisvaikeustutkimuksessa
Diss.

This study addressed the issues of using clinical archival data for research with a sample of children with learning disabilities who were referred for neuropsychological evaluation. Clinical samples are not representative samples, and this must be acknowledged when conducting research. One bias is the increased rate of comorbidity among clinical samples, and they can be used to study the mechanisms of comorbidity. When testing the hypothesis of attention deficits in comorbid reading disability (RD) and attention deficit hyperactivity disorder (ADHD) being secondary to RD (phenocopy hypothesis), it was found that comorbid RD/ADHD group showed deficits in abilities related to both RD and ADHD. The results were not supportive for the phenocopy hypothesis. One frequent problem in clinical datasets is missing data. The study tested several data imputation techniques, their accuracy in predicting missing values, and their subsequent effects on the parameter estimates. Clinical neuropsychological datasets consist of measures covering all domains of neuropsychological functioning and are especially suitable for investigating the properties of neuropsychological measures. Rapid Serial Naming (RSN) tasks require to speedily name serially presented familiar visual stimuli. RSN tasks comprising different stimuli were found to measure partly different skills. RSN performance was related to phonological skills, motor dexterity, verbal fluency, and processing speed. The Trail Making Test-Part B (TMT-B), requires to speedily follow and alternate between numerical and alphabetical sequence. Children with RD were slower than controls on TMT-B. These differences were accounted for by differences in following the alphabetical sequence. Clinical assessment and research differ in several ways, resulting in problems and limitations, but also possibilities, to the application of clinically collected data in research.

Keywords: learning disability, neuropsychology, clinical data, missing data, data imputation, comorbidity, reading disability, attention deficit hyperactivity disorder, Trail Making Test, rapid serial naming

Author's address Vesa Närhi
Niilo Mäki Institute
P.O. Box 35
FIN-40351 Jyväskylä, Finland
E-mail: vesa.narhi@nmi.jyu.fi

Supervisors Professor Heikki Lyytinen
Department of Psychology
University of Jyväskylä, Jyväskylä, Finland

Professor Timo Ahonen
Department of Psychology
University of Jyväskylä, Jyväskylä, Finland

Docent Tapio Korhonen
University of Turku,
Turku, Finland

Reviewers Professor George Hynd
University of Georgia, USA

Docent Pirkko Nieminen
University of Tampere,
Tampere, Finland

Opponent Professor George Hynd
University of Georgia, USA

ACKNOWLEDGEMENTS

There are several individuals who have been involved in and supported me in this project. First, I would like to thank my supervisors. Prof. Heikki Lyytinen, guided and supported me throughout this work. I especially appreciate his ability to balance between giving guidance and responsibility. Prof. Timo Ahonen worked closely with me, his wisdom and support carried me over several difficult phases during these years. Docent Tapio T. Korhonen advised and supported me in this project. My gratitude goes also to reviewers of my thesis, Prof. George Hynd and Docent Pirkko Nieminen, for their encouragement and criticism.

I am also grateful to co-authors of the individual papers. Pekka Räsänen, Riitta-Leena Metsäpelto, Seppo Laaksonen, Risto Hietala, Mikko Aro, Taisto Leppäsaari, and Asko Tolvanen offered their time, effort and expertise in various phases of the studies. My thanks go also to Stephen Lord and Jane Erskine for revising my English.

This work would not have been possible without the persons working at the Niilo Mäki Institute Child neuropsychological clinic. I thank them for their patience in data collection and support when preparing the studies.

This project has been financially supported by The Academy of Finland, Foundation for the Haukkala's Child Psychiatric Institute, and Niilo Mäki Foundation. It has been prepared in collaboration with the project "Human Development and Its Risk Factors".

Finally, I want to thank my family, Tanja, Fiina and Aapo, without whom this work might had been finished earlier, but the time it took surely had been less enjoyable.

LIST OF PUBLICATIONS

Study 1
Närhi, V. & Ahonen, T. (1995). Reading Disability with or without Attention Deficit Hyperactivity Disorder: Do attentional problems make a difference? Developmental Neuropsychology, 11, pp. 337-349.

Study 2
Närhi, V., Räsänen, P., Metsäpelto, R.-L., & Ahonen, T. (1997). Trail Making Test in assessing children with reading disabilities: A test of executive functions or content information. Perceptual and Motor Skills, 84, pp. 1355-1362.

Study 3
Närhi, V., Laaksonen, S., Hietala, R., Ahonen, T., & Lyytinen, H. (2001). Treating Missing Data in a Clinical Neuropsychological Dataset—Data Imputation. The Clinical Neuropsychologist, 15, pp. 380-392.

Study 4
Närhi, V., Ahonen, T., Aro, M., Leppäsaari, T., Tolvanen, A., Korhonen, T.T., & Lyytinen, H. (submitted). Rapid serial naming: Relations between different stimuli and neuropsychological factors.

CONTENTS

1 GENERAL INTRODUCTION

The interplay between clinical assessment and research has a long-standing tradition in child neuropsychology and in the studies of the neuropsychological basis of learning disabilities. Combining clinical work and research has been the frequently selected strategy for many pioneers of developmental neuropsychology. This strategy has greatly advanced the knowledge on developmental neuropsychology and learning disabilities. For many researchers, this strategy has led to the combining of clinical assessment and research data collection.

One example of combining neuropsychological assessment and research is a long series of studies concerning the validity and properties of neuropsychological measures carried out by Reitan and his colleagues. Starting with the development of the Halstead-Reitan Neuropsychological Test Battery for Older Children (HRNTB) in 1955, Reitan and his colleagues adopted a strategy of gaining clinical experience of individual cases before proceeding to group studies on the measures within the test battery. The extensive sequelae of studies pertaining to the use of HRNTB that followed are reviewed by Reitan and Wolfson (1992). The studies addressed the ability of the measures to differentiate children with documented central nervous system damage from control children (e.g. Reed, Reitan & Klove, 1965), the age-related changes in neuropsychological performances (e.g. Reitan, 1971), the development of screening indices for the identification of suspected cerebral dysfunction in children (Reitan & Herring, 1985), and the identification of those functions especially vulnerable to central nervous system damage (e.g. Boll & Reitan, 1972; Nici & Reitan, 1987; Reitan & Wolfson, 1988).

Data collected in the course of clinical evaluation has also been used in the studies of subtypes of children with learning disabilities. Rourke and his colleagues have used large clinical databases in studies concerning the syndrome of non-verbal learning disability (NLD) (Rourke, 1989). These studies have highlighted the importance of recognising the differences among children with learning disabilities, and especially in appreciating the effects exerted by the underlying deficits on the psychosocial functioning of some children with

learning disabilities (Fuerst, Fisk & Rourke, 1989; Fuerst & Rourke, 1993; Rourke & Fuerst, 1992). Also, Rourke and his associates have taken advantage of large clinical data sets in studies relating to the properties of the measures used in neuropsychological assessment (e.g. Brown, Rourke & Cicchetti, 1989; Fisk & Rourke, 1987; Francis, Fletcher & Rourke, 1988).

Another line of study conducted with clinical data sets is longitudinal studies. As the clinical datasets consist of children who have significant difficulties in meeting with academic expectations, they provide a good starting point for studies on the long-term outcome of children with LDs. Spreen and his colleagues have implemented this approach in studies concerning the stability of learning disability subtypes (Spreen & Haaf, 1996), and with regard to the stability of neuropsychological deficits in different subgroups of subjects with LDs (Sarazin & Spreen, 1986).

The combination of clinical assessment and research data collection has proven to be a fruitful tradition in the neuropsychological investigation of LDs. Collecting data within the clinical setting has several advantages over other means of data collection. Some advantages are practical—one does not need laborious and expensive screening to locate children with LDs, data collection can occur, at least in part, as a "side project" of clinical assessment, and data collection can occur on a continuous basis. Another reason for the close relationship between clinical work and research is that many of the research questions are closely related to the questions faced by the clinician in day-to-day practice. All of the above-mentioned research traditions are directly relevant to clinical work. The development and study of assessment methods is an essential prerequisite that allows clinicians to select proper measures with which to evaluate the children. Subtyping research gives clues and concepts for diagnoses and the planning of remediation planning. Longitudinal research gives background knowledge with regard to communicating with children and their parents, especially the latter's worries about the future of their child. Concerning these factors, the combination of clinical assessment and research data collection seems to be a very sensible decision.

However, combining clinical assessment and research data collection is not unproblematic, and the use of such data has limitations. Clinical assessment differs from pure research data collection in several areas, all of which must be taken into account when using clinical archives in research. Collecting research data is, in principle, a straightforward procedure. The research questions are formed, the target populations are identified, the concepts of interest are determined and operationalised, and all the selected measures are obtained from the selected sample. The focus of research data collection is, in terms of questions and hypotheses, formed at the level of the theoretical concepts, and in terms of the interpretation of the results, at the level of the sample.

The focus of clinical assessment is very different from that of research data collection. The possibility to select individuals in clinical practice is limited—at best, exclusionary criteria can be used. The initial questions set for the assessment are not formed by the clinician, but are given from the referring

source, which then naturally affects to the selection of the measures used in the assessment. The focus of the clinical assessment is both, in terms of the questions and hypotheses formed, and in terms of the interpretation of the results, at the level of the individual.

The problems and the possibilities of using data from clinical archives in learning disability research arise largely from the differences in the aims and in the focus between clinical assessment and research data collection. First, as the studied participants cannot be properly sampled from a larger population, clinical samples are always somewhat biased. There are several issues pertaining to referral bias, and in many cases, the exact nature of these biasing issues is not known. One well-known example of a common referral bias is the over-representation of males over females in clinical samples. In clinical samples of children with RD, the boy/girl ratio has been reported to be from 3,5 to 4,0, in general population samples, the ratio has been found to be 1,5 to 1,8 (DeFries, 1989), and in some studies, the prevalence of RD has been found to be equal among boys and girls (Shaywitz, Shaywitz, Fletcher & Escobar, 1990). Likewise, the observed male/female ratios of children with ADHD differ between clinical and community samples (Gaub & Carlson, 1997). Children referred to a clinic for the assessment of learning difficulties are not representative of a larger population of children with learning difficulties, and the ways in which the sample is non-representative are not fully known.

Second, the clinical assessment procedure places limitations on the choice of usable measures. Research data collection can be targeted to highly specific measures of the concepts under study; it is also possible to use a large set of measures within a very limited area of interest. Many research designs also permit the use of experimental measures created for that particular study. As the questions for the clinical assessment are set at the level of an individual, these practices are often very difficult, or even impossible to apply. In order to furnish meaningful answers to the questions posed, the clinical assessment has to cover a reasonably wide range of abilities. Naturally, it is possible, and often necessary, to assess some limited skill areas in greater detail, but the important difference is that the areas of more detailed assessment are determined separately for each individual subject, not by the pre-set research questions. The same holds true for the use of experimental, new measures. For these reasons, the data on the specific measures are likely to be available from a very selected population in the clinical archives, and consequently, the possibility to use them in research is limited. Another practical limitation is that since clinical practice in an ongoing procedure, it is very difficult to use measures that require continuous access to special settings, for example to laboratory. The data in the clinical archives that is available from reasonably large samples are likely to be based on standard neuropsychological measures.

A third common problem arising from the different aims of clinical assessment and research data collection is missing data in the clinical archives. As with clinical assessment, the objective is to provide relevant information in terms of the problems faced by the child within a limited time. Thus, the

clinician is frequently placed in the situation where he/she has to choose between obtaining complete data on all measures planned to be included in a dataset, and obtaining the most relevant data in terms of the clinical assessment. Naturally, the clinical needs are pre-determined by ethical reasons. The outcome of this frequently results in missing test items from a large portion of the sample, and these missing items pose problems for the use of the dataset for research purposes.

1.1 Sampling and Comorbidity

The possibility of obtaining non-representative samples when studying clinical data is evident. In all research concerning LDs, some selection of the participants is required. When studying clinical populations, the first screening is always instigated by the referrer. In the case of, for example, RD, there are several factors other than difficulties in reading which may result in referral (and thus also run the possibility of being included in the research sample). These other factors can be related to social policy matters (e.g. services available in certain areas), to the expertise of the referrer (e.g., one trained and competent in reading instruction may be less likely to refer the child with RD than another without such competence), to the variety and severity of the difficulties related to the child's LD (e.g. disagreement between parents and school personnel on the actions that need to be taken), and to the severity and multiplicity of the child's learning problems.

There is evidence for referral biases concerning both children with reading difficulties and with hyperactivity problems. Shaywitz, Shaywitz, Fletcher and Escobar (1990) used two methods, school identification, and discrepancy between general intelligence and reading achievement, to identify children with RD in a same population, and studied the identification method dependent differences. Two main findings emerged. First, there was clear over-representation of males in the school-identified group relative to the discrepancy criteria identified group. Based on this finding they suggest that the reported increased prevalence of RD in boys may be a result of referral bias. The second main finding related to the problems other than reading skills of the identified children. Children with discrepancy criteria identified RD, but without behaviour problems, were under-represented in the school identified group, while children without RD but with behaviour problems according to the discrepancy criteria, were over-represented in the school identified group. It seem clear therefore that factors other than reading skills influence teachers' perceptions of the child, and consequently, the decisions regarding the identification of the child as having, or not having RD.

Woodward, Dowdney and Taylor (1997) compared clinic-referred and non-referred children with hyperactivity problems. While the two groups did not differ in the severity of hyperactivity, there were several differences

between the groups. The clinic-referred children had significantly more conduct disorders and emotional disturbances in comparison to non-referred children. The parents of the referred children exhibited more negative parenting behaviours and were more depressed than the parents of non-referred children. The higher incidence of mixed disorders in the referred sample is in accordance with the higher than expected rate of comorbidity in the clinical samples. The results of both Shaywitz et. al. (1990) and Woodward, Dowdney and Taylor point to the fact that the clinical referral of a child is a result of complex interplay between the different characteristics of the child and socio-familial characteristics.

There is evidence that children affected by the same disorder differ in important ways in different clinical populations. Epstein, Shaywitz, Shaywitz and Woolston (1991) studied children with ADHD who were seen in different clinical settings. The children from different clinics were similar in terms of severity of attention problems and hyperactivity, but there were source dependent differences in the amount of learning disabilities and behavioural and psychiatric problems incurred by the children. Clearly, had the associated problems of ADHD been under study, then the results would have varied depending on the clinic from which the study population was recruited. Some of these matters affecting the possibility of referral are very difficult to define, and the possible effects exerted by the referral bias of a particular clinic on the results when using data from that clinic in research is, in many cases, impossible to evaluate. The point to keep in mind is that when using clinical data in research, one should be aware of the existing biases in the sample, and consequently, be cautious when generalising the results.

One particularly important bias in clinical populations is the increased rate of comorbidity. It is well acknowledged that school related disorders are more likely to co-occur than could be expected by chance. Clinical samples will always show, irrespective of referral bias, an increased rate of comorbidity (Caron & Rutter, 1991). Comorbidity at the group level has to be taken into account when interpreting the study results—one can not assume that children expressing one particular disorder in the clinical sample are representative of all children with that disorder. Comorbidity can severely obscure the results of studies targeted at factors related to a specific disorder, and thus there is always the possibility that the results obtained are a consequence of another occurring condition which is over-represented in the sample. The increased rate of comorbidity in all clinical samples has to be taken into account when studying the factors related to a particular disorder, and the results obtained should be interpreted with caution, and if possible, controlled for the comorbid conditions.

Caron and Rutter (1991) have addressed the issue of comorbidity in child psychopathology. Because of the bias towards high rates of comorbidity clinical datasets cannot be used when studying the occurrence of comorbidity of different disorders. Caron and Rutter have conceptualised the possible causes of observed comorbidity. Increased rates of comorbidity in clinical samples can

result in uncertainty of the results obtained with such datasets. However, clinical datasets offer a rich source of participants with comorbid conditions, making it easy to obtain comorbid groups of sufficient sample sizes. These can be used (acknowledging their limitations), to study the mechanisms of comorbidity (see also Angold, Costello & Erkanli, 1999).

Within the population of children with LDs and other school-related problems, RD and ADHD are the most common disorders. For both RD and ADHD, the estimated prevalence is 5% of the school age population (American Psychiatric Association, 1994; Shaywitz, Shaywitz, Fletcher & Escobar, 1990). The expected comorbidity of RD and ADHD is thus 5% * 5% =.25%. However, the true comorbidity of RD and ADHD is much higher (e.g. Dykman & Ackerman, 1991; Shaywitz, Fletcher & Shaywitz, 1995). There are several possible reasons for this higher than expected rate of comorbidity. One can assume that one of the disorders is a cause of another—RD may cause ADHD-type behaviour as a secondary disorder due to frequent academic failures; ADHD may cause RD as a secondary disorder due to difficulties in attending to reading instruction. It may also be that both of these disorders have a common underlying deficit, which, together with other, e.g. environmental, influences, may cause RD in some children; ADHD for others, and both RD and ADHD for another group of children. It is also possible that children with both RD and ADHD differ from both of the 'pure' groups.

1.2 Neuropsychological measures

To plan and compile a neuropsychological assessment battery is a theory-driven endeavour, and the resultant assessment battery can be seen as an operational definition of the theoretical model. The aims are to construct measures of different components of the theoretical model, and to include measures of all components in the assessment. The relation between theory and assessment differentiates neuropsychological assessment batteries from most intelligence measures used with children such as WISC-III (Wechsler, 1991), the aim of which is to give a large scale picture of the cognitive abilities of the individual, without specific regard to theories of the structure of neuropsychological or cognitive abilities.

The theoretical background of the neuropsychological assessment batteries is clear. For example, the development of the HRNTB is based on the Reitan-Wolfson model of neuropsychological functioning (Reitan and Wolfson, 1992). The model includes sensory and motor functions and three levels of central processing, and each of these is covered within the assessment battery. Another example of the theory-assessment relationship is the development of the NEPSY (Korkman, Kirk & Kemp, 1998), which is based on the Luria's (e.g. 1980) theory of the functional units of the central nervous system. The purpose of the clinical neuropsychological assessment is to cover all areas of

neuropsychological functioning, and to interpret each test result in relation to other results (Reitan and Wolfson, 1992).

The aim of the neuropsychological assessment is to use measures that are related to specific concepts, and are in this sense 'pure'. In practice, it is difficult to achieve, and most of the tests used are multi-componential, i.e. the performance is determined by several different abilities. Just looking at a specific test score the question remains as to 'what is the precise reason that the child performed poorly/well on the test?', and often, there are several possible explanations. This question can be dealt with effectively within the clinical setting where it is possible to test different hypotheses by using the rich amount of information gathered. This can be achieved by relating the performances on different tests of the battery to each other, by using additional assessment methods if required, and by interviewing the child, parents, and teachers. The hypotheses are formed and additional data obtained individually for each client. The result of this is that in the archives there is additional data to aid the interpretation of each measure for some, but not for all subjects. Also, a large part of the information is obtained from non-structured interviews and thus is difficult to quantify.

From the point of group level research, the usable data in the clinical archives consist of neuropsychological measures that are multi-componential, and consequently, the interpretation of which is difficult. However, the clinical archives also provide ways to overcome the problem of multi-componentiality. As described, the aim of clinical neuropsychological assessment is to cover all domains of neuropsychological functioning. Since different abilities affect differently on performances on different tests, the multi-componentiality of the tests can be evaluated or controlled for by using information regarding the relationships between different measures within the test battery. Thus, this inherent quality of clinical assessment can be used to advantage when using clinical data for research, and to clarify the properties of neuropsychological measures by studying performance on a measure by connecting it to performances on other measures. This kind of approach will shed light on the abilities reflected by different measures and will improve the validity of the use of these measures, both in clinical practice and in research. A notable example of this kind of research, in addition to the above-mentioned work of Reitan and Rourke and their colleagues, are the studies by Snow (e.g. Snow, 1998a, 1998b; Snow, English & Lange, 1992).

The multi-componentiality is the most apparent problem when assessing executive functions (EFs). Since, by definition, EFs are higher order controlling functions, their assessment requires the use of materials that are also sensitive to more basic functions. Consequently, poor performance on tasks purported to measure EFs can reflect deficits in EFs as well as in basic functions operationalised in the task. Denckla (1994, 1996) has especially addressed this problem whereby she proposes that the lower level skills should be systematically controlled for when measuring EFs.

The multi-componentiality of measures in the area of EFs becomes a question of a test being a valid measure of EFs. The multi-componentiality in some other cases becomes a theoretical question—how should it be explained that performance on a particular test is related to a particular outcome measure? The questions then are, for example, 'what skills affect performance on the test?' and, 'which skills are critical to the explanation of the relationship between test performance and achievement or behavioural outcome measure?'. As the clinical assessment covers a wide range of abilities, clinical data sets are helpful in answering questions related to the properties of assessment methods.

1.3 Missing data in the dataset

The nature of the clinical assessment very easily results in use of measures that differ according to the needs of the individual client. Clinical assessment always occurs within given time limits, and the purpose of the clinical assessment is to provide diagnostic information relevant to the alleviation of the child's problems. When trying to obtain relevant answers it is, in many cases, more relevant to assess other skills than those covered by tests planned for research purposes. In many cases, it is not possible to meet both the aims of clinical assessment and data collection, and when this happens, the clinical needs are naturally the priority. The result of this is missing test items in the archives, and from the research perspective, missing data in the dataset.

The problems caused by missing data vary naturally from one research design to another. In many cases with clinical datasets it is relatively easy to conduct group comparisons on a small number of measures, or case studies, where all necessary information can be quite easily obtained. When it comes to designs that require a large number of participants and a large number of measures, missing items may seriously affect the utility of clinical data in research. This is especially the case when using multivariate methods of analyses, where the main impact is the reduction of statistical power and increased uncertainty of parameter estimates.

The effects exerted by the missing data on the parameter estimates depends on the mechanisms by which the data are missing. In the clinical setting, the data are not likely to be missing completely at random (i.e. occurrence of missing data not depending on any factors inside or outside of the data), but what is missing depends on several non-random factors such as the referral question and purpose of the assessment, and on the nature of the client's problems. In any case, when the data are not missing completely at random, the obtained parameter estimates will be biased—this being almost always the case in clinical datasets.

In many of the statistical methods, a general presumption is complete data. If this is met by excluding all cases with missing values, the result is a reduction in the number of participants in the analysis. In some cases this

makes analysis impossible, in any case, the result is a reduction of the statistical power of the analysis. Furthermore, if the data are not missing completely at random, this deletion of the cases functions as an additional screening phase for the subjects, and the remaining sample may not be representative of the original sample (Schafer, 1997).

Data imputation refers to a technique whereby auxiliary information is used to predict missing values and these predictions are used to complete the dataset. The theory of data imputation has advanced rapidly, but imputation methods have not previously been applied to complete clinical neuropsychological datasets. As clinical datasets hold a large amount of information from each participant, this information can be used in imputation models to predict the values of missing observations. Testing and developing data imputation techniques in clinical datasets enhances their use in research and helps to combine clinical assessment and statistical decision making.

2 AIMS OF THE EMPIRICAL STUDIES

The purpose of this dissertation was to address the problems and possibilities resulting from the combining of clinical neuropsychological assessment and scientific research based on a quantitative rationale. The aim of Study 3 was to address the problem of missing items in the dataset. Advanced theory of data imputation has developed rapidly, but applications of imputations to neuropsychological datasets have been rare. Since the clinical dataset generally has additional data to that selected for research purposes, it can be used as a basis for imputation modelling. Consequently the possibilities of applying imputation to these datasets are evident. Since there were no previous publications on the application of data imputation in clinical child neuropsychological datasets, we tested several imputation methods before selecting one to be used with our dataset.

Two studies (Study 2 and Study 4) address the issue of the multi-componentiality of neuropsychological methods. For both of the studies an additional sample from the general school population was used. In Study 2, an experimental measure was designed and used to study the validity of one commonly used measure of EFs in a population of children with RD. In Study 4, the rich quantity of neuropsychological test data obtained during clinical evaluation was utilised to study the factors relating to rapid serial naming performance.

Study 1 was targeted at the examination of the comorbidity between RD and ADHD. Comorbidity is more common in clinical samples than in epidemiological samples, and if this issue is not addressed, complications arise in the interpretation of studies concerning learning disabilities. Study 1 was conducted to directly test the hypothesis relating to the comorbidity of RD and ADHD presented by Pennington, Groisser and Welsh (1993).

3 METHODS

The sample used in the studies was obtained from the archives of the Niilo Mäki Institute Child neuropsychological clinic (NMI-clinic). The NMI-clinic is specialised in the assessment of children with learning disabilities.

3.1 Referral procedure of the children

The clinic serves the area of Central Finland as part of the Family Guidance Services. The children were referred to the clinic following psychological evaluation. As the clinic is community based, no formal exclusionary or inclusionary criteria were applied to the referral procedure. The instructions for the psychologists working at the sources of referral included that the main problems resulting to the original referral of the child should be learning disabilities and/or attention deficits, with IQ within normal range. If the conclusion on initial evaluation was that the problems were likely to be due mainly to the family situation or socio-emotional factors, referral to the NMI-clinic was not recommended. As a result of these instructions, most children whose main problems were socio-emotional, as well as those children suffering from a more general developmental delay were not referred to the NMI-clinic. No specific age limits were applied, although it was recognised that the assessment at the NMI-clinic was most appropriate for children from grades 1 to 6 (aged 7 to 13 years). The most common referral source was Family Guidance Clinics, a minority of children were referred by psychologists working at primary heath care centres, and by the Central Hospital of Central Finland.

3.2 Original archive data

The original and complete dataset included all children assessed at the clinic between the years 1985 and 1997. The first two studies (Studies 1 and 2, since they were conducted prior to the termination of data collection) utilised a sub-sample available for each study at that period of time.

From the outset, the aim of the clinic was to combine clinical work and data collection for group studies. To achieve this goal, a set of neuropsychological measures was selected for administration to all children presenting for assessment at the clinic. The measures were selected to comprehensively cover areas of neuropsychological functioning, especially those of main importance in the assessment of children with suspected LDs. The measures are presented in Table 1. In addition to the neuropsychological measures, child behaviour rating scales were used (CBCL, Achenbach & Edelbrock, 1986a, and Achenbach, 1991a; CBCL-TRF, Achenbach & Edelbrock, 1986b, and Achenbach, 1991b), as well as achievement tests on reading and arithmetic (Niilo Mäki Institute, 1992). All children were intended for assessment with this standard test battery and in addition, complementary measures were used to assess the skills relevant to each individual child.

The aim of the data collection was to obtain similar information for all children visiting the clinic, and thus to cover information pertaining to all main domains of neuropsychological functioning relevant to the assessment of children with LDs and attention deficits. This aim was not completely achieved, due to changes in the measures deployed during the clinic's history. As the data collection period spanned several years and more and more relevant research on the neuropsychology of LDs emerged, it was necessary to include new measures proven to be valuable in the assessment of children with LDs. Also, some measures, which were of potential theoretical importance, but according to clinical experience were found to be uninformative for the majority of the assessed children, were excluded from the assessment battery, or alternatively, were administered only to those individual children whose assessment was thought to especially benefit from the information yielded by the measure. Naturally, had there been unlimited time for the assessment, all original measures could have been administered to all children, but unfortunately this was not the case.

The necessary changes in the assessment methods resulted in the fact that only a portion of the original measures were given to a sufficient number of children to enable their use in any larger scale study. In this sense, the original plan of collecting complete data on all referred children was only partly successful. On the other hand, as the focus of the clinical evaluation is on the individual child, the changes were of great importance to the provision of the most relevant information on each child. These kinds of changes in the assessment methods are unavoidable in any clinical environment purporting to long-range research goals with data collection periods spanning several years.

In Studies 1 and 2, only a sub-sample of children and measures were utilised. In Study 1, which was conducted prior to the completion of the whole data collection, the participants were selected on the basis of the available test data. In Study 2, an experimental measure was created and implemented with all consecutive referrals starting from the design of the study, and to which all children fulfilling the inclusion criteria of the study were selected.

TABLE 1 Neuropsychological measures originally included in the NMI-clinic test battery

General Intelligence	
*Wechsler Intelligence Scale for Children—Revised (Wechsler, 1974)	
Handedness and tactile perception	
*Edinburgh Handedness Test (Oldfield, 1971)	*Tactile Finger Recognition (Reitan & Davidson, 1974)
Left-Right discrimination test (Benton 1959)	Fingertip writing (Reitan & Davidson 1974)
Tactile Form Recognition (Reitan & Davidson 1974)	
Verbal and auditory abilities	
*Peabody Picture Vocabulary Test (Dunn & Dunn, 1981)	*Verbal fluency (Benton, Hamsher & Shian, 1994)
*Boston Naming Test (Kaplan, Goodglass & Weintraub, 1983)	*ITPA Sound Blending (Kirk, McCarthy & Kirk, 1968; Finnish version, Kuusinen & Blåfield, 1972)
*Token test (shortened version, DeRenzi & Faglioni, 1978)	*Rapid Automatized Naming (Denckla & Rudel, 1976; Wolf, 1986)
*Morfology Test (Lyytinen, 1987)	*Nonword reading (Niilo Mäki Institute, 1992).
Seashore rhythm test (Knights & Norwood, 1979, Reitan & Davidson, 1974	
Visual and visuo-motor skills	
*Gestalt Closure (Kaufman, 1983)	*Trail Making Test (Reitan & Wolfson, 1985)
*Developmental Test of Visual-Motor Integration (Beery, 1989)	Judgment of line orientation (Benton, Hannay & Varney, 1975; Benton, Hamsher, Varney & Spreen, 1975)
Target Test (Reitan & Davidson, 1974)	Face recognition (Benton & Van Allen, 1968; Benton et al., 1975)
Problem solving and executive functions	
*Coloured progressive matrices (Raven, 1965)	*Wisconsin Card Sorting Test (Heaton, 1981)
Memory	
*Selective Reminding Test (Buschke & Fuld, 1974)	*Spatial memory (Kaufman, 1983)
*Digit Span (Wechsler, 1974)	Recurring Figures (Kimura, 1963)
Word Order (Kaufman, 1983)	Nonverbal Selective reminding test (Buschke & Fuld, 1974)
Sentence repetition task	Story repetition (Korkman, 1980)
Motor skills	
*Purdue Pegboard (Tiffin, 1968)	Foot tapping (Knights & Moule, 1968)
Finger tapping (Reitan & Davidson, 1974)	Grip strength (Reitan & Davidson, 1974)
Maze test (Knights & Moule, 1968)	

Note. Measures marked with asterice were included in the final dataset

The whole applicable sample was used in Study 3. As the aim of the study was to test and choose an imputation method used to be able to complete the data

set, all children and all measures were considered for potential inclusion in the study. As the clinic used no strict inclusion criteria, there was a need to set specific criteria for the children to be included in the dataset from all children assessed thus far. The first inclusion round was conducted by selecting participants to the dataset. The age range was set at 8 to 11 years. This was because the entire assessed sample included a very limited number of younger and older children, and because most of the assessment methods used were most appropriate for this age range. Other exclusionary criteria were no uncorrected sensory handicaps, no severe emotional disorders, Finnish as a native language, no acquired central nervous system damage, no excessive absence from school, and a general intelligence level within normal limits. Also, as the main interests focus on reading disabilities, the children had to have been administered with a reading achievement measure as part of the evaluation.

The second selection round was conducted to select the neuropsychological measures for the dataset. The foci here were the amount of missing data and the coverage of neuropsychological functions. With the criterion of no more than 15% of missing data points, 20 tests were selected to be included in the dataset (Table 1, marked with asterice). A 'missingness' rate of 15% represents a reasonable amount of missing items in terms of data imputation (Roth, 1994). With this imposed limit, neuropsychological functions were also covered with an acceptable level of comprehensiveness. The third selection round was conducted again on participants whereby those children fulfilling the criteria imposed at the first round were allowed to have missing data on five but on no more of the 20 variables. Of the 361 children in the data set, 193 fulfilled these criteria, and the missing items were imputed as described in Study 3, and this imputed dataset was used in Study 4.

4 THE ORIGINAL STUDIES

Study 1

Reading Disability with or without Attention Deficit Hyperactivity Disorder: Do attentional problems make a difference?

by

Vesa Närhi and Timo Ahonen
Developmental Neuropsychology, 1995, 11, 337-349.

https://doi.org/10.1080/87565649509540624

Study 2

Trail Making Test in assessing children with reading disabilities: A test of executive functions or content information.

by

Vesa Närhi, Pekka Räsänen, Riitta-Leena Metsäpelto and Timo Ahonen
Perceptual and Motor Skills, 1997, 84, 1355-1362.

https://doi.org/10.2466%2Fpms.1997.84.3c.1355

Study 3

Treating Missing Data in a Clinical Neuropsychological Dataset—Data Imputation.

by

Vesa Närhi, Seppo Laaksonen, Risto Hietala, Timo Ahonen and Heikki Lyytinen
The Clinical Neuropsychologist, 2001, 15, 380-392.

https://doi.org/10.1076/clin.15.3.380.10266

Study 4

Rapid serial naming: Relations between different stimuli and neuropsychological factors.

by

Vesa Närhi, Timo Ahonen, Mikko Aro, Taisto Leppäsaari, Asko Tolvanen, Tapio T. Korhonen and Heikki Lyytinen

Manuscript (submitted) 2001

https://doi.org/10.1016/j.bandl.2004.05.004

5 GENERAL DISCUSSION

The aim of the dissertation was to address the problems and possibilities posed by the use of clinically-collected neuropsychological data in the research of learning disabilities. The studies addressed problems caused by missing data in the archives and the ways in which to develop imputation methods for such datasets, the multi-componentiality of neuropsychological measures, and the comorbidity of developmental disorders.

5.1 Sampling and comorbidity

The starting point in research that utilises clinically-collected data is that there always biases in the sample. There are various issues affecting these biases, these issues are likely to vary from one clinic to another, and, for the most part, all biasing factors cannot be identified. Bias in the sample ensures that the results obtained may not be generalisable to a larger population. Because of the inherited quality of the clinical dataset, one should be careful with the conclusions drawn from studies using such datasets and thus replications with both other clinical and general population samples are of great importance.

One factor on which clinical datasets are known to be biased, is comorbidity. Within the area of LDs and other school-related problems, comorbidity is well established in general population samples. Clinical samples, in which there are a large number of comorbid participants, provide an invaluable source for studies targeted at the mechanisms of comorbidity.

Study 1 addressed the comorbidity of RD and ADHD, testing the hypothesis presented in another study that utilised a clinical sample (Pennington, Groisser and Welsh, 1993). Based on their results, Pennington et. al. concluded that the attention-hyperactivity problems of children with reading disability are secondary, and that comorbid children do not show neuropsychological deficits characteristic to children with pure ADHD, but instead show deficits characteristic only of children with RD. The findings of

Pennington et. al. concerning the neuropsychological characteristic deficits of a comorbid RD/ADHD group were found not to apply to the present sample. There are a number of studies targeted at the same issue, most of which have not provided support for the 'phenocopy hypothesis' of Pennington, Groisser and Welsh (1993) (e.g. Nigg, Hinshaw, Carte & Treuting, 1998; Reader, Harris, Schuerholz & Denckla, 1994; but see Shaywitz, Fletcher & Shaywitz, 1995). The research concerning the neuropsychological deficits associated with comorbid RD and ADHD highlights the importance of replications.

The issue of the comorbidity of RD and ADHD is as yet unsolved. There may be several mechanisms responsible for the increased rate of comorbidity. Several studies suggest a common aetiology for both RD and ADHD, for at least a subset of the children having both disorders. In genetic studies, heritable variation has been found to account for a significant amount of the covariance between reading and hyperactivity (Light, Pennington, Gilger & DeFries, 1995), and that the comorbidity of hyperactivity and spelling disability is largely due to common genetic influences in the two (Stevenson, Pennington, Gilger, DeFries & Gillis, 1993). It is also possible that there are different subgroups of children with ADHD, whose attention difficulties are due to different dysfunctional attention systems, and that RD is specifically related to only one subgroup of ADHD (Halperin, Newcorn, Koda, Pick, McKay & Knott, 1997). It seems reasonable to suppose that part of the comorbidity is due to both ADHD and RD functioning as an additional risk factor for the other (Schulte, Conners & Osborne, 1999).

Clinical samples are never random samples of participants with particular disorders or deficits. Within the area of LDs and other school-related problems, there are various issues affecting the possibility of a child being referred for clinical evaluation. Optimally and theoretically, these factors could be defined, isolated, and subjected to study, in practical situations that is very difficult, if not impossible. In practice, the best that can be achieved is to acknowledge these factors, and try to identify them and interpret the results accordingly. Some questions, such as the neuropsychological deficits associated with comorbid RD and ADHD, can be examined with different populations and with different measures, and reliable answers obtained through replications.

5.2 Missing data

The aims of the clinical assessment and research data collection are different, and within the clinical setting, it is natural that the needs arising from the clinical assessment are paramount with the requirements of research data collection being met if possible. One resultant factor is that not all children are administered with all of the measures planned for inclusion in the research dataset, and, from the perspective of research data, there are missing test items. Incomplete data can severely affect the utility of data, particularly in research

utilising multivariate statistics, and, in some cases, make the statistical methods inapplicable to the data. Data imputation refers to various techniques in which missing items are predicted by the use of modelling. The theory underlying data imputation has developed rapidly, but it has not been applied to clinical neuropsychological datasets.

The aim of the NMI clinic was to combine clinical evaluation and research data collection in a systematic manner, and obtain the same information for all the children visiting the clinic. The clinical needs were considered as primary, and together with more random reasons, this resulted in incomplete data for most of the children under study. Thus far, the data have been adequate for group studies involving only a limited number of measures (Studies 1 and 2), as well as in case studies (Lamminmäki, Leppäsaari, Siiskonen, Ahonen & Lyytinen, 1994; Ahonen, Leppäsaari, Lamminmäki, Poikkeus & Siiskonen, 1997). When attempting to apply the data to studies that require a large numbers of participants, measures, and subsequent analysis using multivariate statistics, the missing data caused severe problems. To solve the problem of missing data, data imputation was applied. Since there were no previous applications of data imputation techniques to clinical neuropsychological datasets, several imputation methods were tested on the data.

Subsequent to the screening phases, a sample of sufficient size with a reasonable rate of missing items on measures covering the main areas of neuropsychological functioning was identified. Four different imputation models were placed under close scrutiny (after discarding the most simple methods due to their exerting severe effects on the parameter estimates) by artificially deleting some information. The accuracy with which different methods were able to predict the actually obtained values and the changes that they produced on the parameter estimates were studied. Of the models, the one with the most acceptable levels of both accuracy and changes produced on the parameter estimates was selected for use when completing the dataset.

The selected procedure followed the logic of clinical hypothesis formation—the critical variables of the model included the child's age, reading disability status and the two WISC-R sub tests yielding the highest correlations with the neuropsychological measure to be imputed. This procedure performed better than other, more purely mathematically-defined imputation models. It appears that utilising the knowledge on the subject matter aided the development of an imputation model well-suited to the present dataset.

An analysis of the underlying reasons for missing observations is the starting point for data imputation. In the NMI-clinic dataset, the main reasons for values being missing were related to the choices made by the clinician (the need to assess some other areas more closely than others or omitting a test due to time constraints). Naturally, the characteristics of the children also affected the probability of certain measures being missing, but only indirectly, by affecting the clinicians' decisions. Refusals and total incapability of performing the test were only very rarely the reasons for missing values. This analysis of the reason for missing observations in the NMI-clinic's data set leads to the

conclusion that in most cases, the tests omitted were those that were most irrelevant in explaining the child's problems. It appears safe to assume that in this kind of situation, the prediction of missing values by imputation modelling is quite accurate.

Naturally, the mechanics of missing data are different in different data sets and they vary from one patient group to another and thus have to be analysed and appreciated accordingly. Smeding and de Koning (2000) analysed the reasons for missing values in neuropsychological measures of patients with frontotemporal dementia. Frontotemporal dementia is progressive and is characterised by personality changes, behavioural disorders, and deterioration of especially language and executive functions. The highest rate of missing values was on measures of executive functions and reasoning capabilities, and their occurrence was related to the severity of behavioural disturbances. Clearly, the missing values in their data were not missing completely at random but were dependent on the neuropsychological domain being assessed and on the progression of the disorder. By utilising the analysis of missing data and taking advantage of the knowledge on frontotemporal dementia, Smeding and de Koning replaced those values missing due to behavioural problems with the lowest observed values. They claim that this procedure resulted in a more differentiated picture of cognitive deficits related to frontotemporal dementia. Based on their observations, they suggest that studies that exclude cases with missing values are studying non-representative sample of patients with frontotemporal dementia.

Another good example concerning the analysis of reasons for missing values, specifically refusals to perform a task, is presented by Mäntynen, Poikkeus, Ahonen, Aro & Korkman (submitted). Analysis of the neuropsychological test performances of 3,5 year old children showed that children who frequently refused to perform a test, when compared to children with only a few test refusals, also had lower scores on several measures that they did perform, scored lower on overall developmental measures, and also had higher refusal rates for assessments at other ages. The results indicate that the test refusals (missing test items) of young children are not random, and when occurring frequently, may be indications of developmental risk.

The problem of missing data is recognised, and consequently, imputation methods are being developed and these are becoming included the statistical packages. For example, new version of LISREL-program (Jöreskog, Sörbom, du Toit & du Toit, 1999) includes full information maximum likelihood method for replacing missing data. The performance of these methods with regard to clinical neuropsychological data is still to be shown, but it seems that the statistical assumptions (normality of the distributions, data missing completely at random) of these models are difficult to meet in clinical data. Furthermore, maximum likelihood algorithms utilise all the information in the data set as a basis for imputation. This can pose a problem relating to further use of the data, resulting in a situation where the same information that has been the basis for imputation is the subject of a further study. The clinical data collected at

different levels (referring symptoms, developmental history, behavioural observations and a wide variety of tests) also make it possible to use information outside of the research data as a basis for imputations (as was a deliberate goal in Study 3). Also, the imputation model developed in Study 3 does not require the data to be missing completely at random. Clearly, there is a lack of studies that address the missing data problem in clinical neuropsychological datasets and consequently, studies addressing the applications of different imputation models to such datasets are required.

5.3 Multi-componentiality of neuropsychological measures

Neuropsychological tests form the core of the clinical evaluation of LD related disorders and thus they also form the core of the clinical data sets. Most of the tests used in neuropsychology are multi-componential—although principally measuring specific neuropsychological functions, the outcome is contingent on several other functions. This multi-componentiality is a problem when interpreting the test results, both in clinical assessment and in research. It is typical to cover a wide range of functions in clinical assessment, and this makes clinical data very useful for the clarification (by studying the interconnections of different measures) of the functions affecting performance in different neuropsychological tests.

Studies 2 and 4 addressed the different abilities affecting performance on two commonly used measures in neuropsychology. However, the underlying questions of the studies differed. Study 2 targeted the validity of using one commonly used measure of executive functions, the Trail Making Test, Part B (TMT-B) with a population on which the test is frequently used. The research questioned the level of performance on different components of the TMT-B for children with and without RD. For the study, an experimental version of the TMT (TMT-A-alphabetic) was constructed and controlled for a previously uncontrolled component of the TMT-B, the speed at which the child could follow the alphabetic sequence. The results showed that controlling for the speed of following the alphabetic sequence is essential to the interpretation of poor TMT-B performance when it is used as an index of an executive function deficit in populations where reading difficulties are common.

The study is a good example of the discourse between clinical and research work. The hypothesis highlighting poor ability in following the alphabetic sequence as a critical factor in TMT-B performance emerged in the clinical work. In turn, this resulted in the construction of the alphabetic version of the TMT and its subsequent use in clinical practice. Instigating control of the speed of following the alphabetic sequence was a good example of clinical hypothesis formation. In the case where a child performed adequately on the TMT-A but poorly on the TMT-B, there were at least two hypotheses; an executive function deficit, or difficulty related to the following of the alphabetic

sequence. These hypotheses guided the further assessment, and led to the creation of the TMT-A-alphabetic. Since this method proved informative in clinical assessment, a more systematic administration of the TMT-A-alphabetic was conducted and the clinical observation was tested empirically with quantitative research.

As Study 2 was targeted at the validity of a measure, the aims of Study 4 were to provide answers to theoretically relevant questions. A deficit in rapid serial naming (RSN) is frequently found among children with RD (e.g. Denckla & Rudel, 1976; Wolf, 1986). The first focus of Study 4 was the connections between RSN measures that differed in terms of the stimulus quality. Differences between these paradigms have previously been hypothesised, but to date, have not been thoroughly tested. The results showed that RSN paradigms that differ in stimulus quality reflect partly different skills.

Different explanations of how deficits in RSN are related to reading performance have been proposed. One explanation is that slow RSN performance of children with RD reflects the frequently observed deficit in phonological skills of these children (e.g. Wagner, Torgesen, Laughon, Simmons & Rashotte, 1993). Others claim that RSN is a far more complex measure than of purely phonological skill, and that slow RSN performance of children with RD is explicable by factors other than phonological skills related to reading performance (e.g. Wolf, Bowers & Biddle, 2000). The studies that have addressed the abilities that affect RSN performance have been conducted with a very limited number of measures, and in no previous study have all the likely relevant components been included (see Wolf, Bowers & Biddle, 2000). With the clinical data set it was possible to include possibly the most relevant components, and to study RSN performance by taking into account the full covariance structure of the explaining variables. The results showed that RSN performance is related to a number of other neuropsychological measures, thus confirming the multi-componentiality of the measure.

Again, as in Study 2, along with the evidence from research, the clinical experiences guided the research. During clinical assessment of children with LDs it had been observed that the child's performance may not be uniform on all RSN measures and that a child can have difficulties only on some of the RSN measures implemented. These observations, together with previous studies, guided the formation of the hypothesis concerning differing qualities among the RSN measures.

The different skills relating to RSN performance were observed within the clinical assessment. Particularly, an attempt to define RSN as a measure of phonological skill has shown to be questionable, since low associations between measures of phonological abilities and RSN are frequently observed. However, some components affecting RSN performance that could not be detected in clinical experience alone, were in this case, made apparent by the research-based analysis of the data. The results of Study 4 further guide the clinical practice by pointing out the relationships between RSN and other measures. Clearly, as discussed in Study 4, some of these observations raised new

questions and hypotheses, relevant to both clinical assessment of the children and for research.

6 CONCLUSIONS

Clinical assessment and research differ in several ways. These differencies pose problems and limitations for use of clinically collected data in research. However, such data also have some advantages. Clinical datasets are especially useful in studying the properties and interrelations of neuropsychological measures. The collection of the large volume of data required for many such studies is a laborious and expensive enterprise if conducted independently of clinical assessment. Such data are routinely collected in clinical assessment.

However, combining clinical assessment and research data collection requires definitive pre-planning. The requirements for the data that is used in research are not similar to the requirements of the clinical evaluation. Furthermore, when working in the clinical setting, there is always the natural risk that the requirements of research data collection must be set aside in order to focus fully on the needs of the individual child in clinical work. In spite of the deliberate philosophy to collect research data at the NMI-clinic, this competition between the clinical and research environments was evident on several occasions. The situation is arguably the norm in most clinical settings. Clear evidence of this competition is shown in the prevalence of missing data on almost all of the measures and for most of the participants. The problem of missing data cannot be overlooked and every good solution to the problem is clearly based on knowledge of the subject matter, starting with the analysis of the reasons for data being missing.

The problems created by biases within clinical data sets have been addressed several times throughout this dissertation. One acknowledged bias in clinical samples is higher than expected rates of co-morbidity and this was the focus of one study in this dissertation. There are several other issues by which the clinical samples are likely to be biased. Unfortunately, other biasing factors could not be put under study here, nor have they been systematically studied elsewhere. The reasons for this are quite clear and practical; biases are likely to vary from one clinic to another and the biasing factors are dependent on several issues and thus are very difficult to define. In any case, when conducting studies with referred clinical populations, the possible non-representativeness

of the sample should always be borne in mind, the results should be generalised with caution, and replications of the studies with different samples should be carried out before any deeper conclusions can be drawn. Naturally, these limitations apply to the studies reported in this dissertation. The requirement for replications is clearly shown with regard to studies of the neuropsychological factors of co-morbid RD and ADHD. On reflection of their original study, and following both a review of other studies reported on the topic and also a replication of their study, Willcutt, Pennington, Boada, et. al. (2001) discuss "...it is possible that the fairly small comorbid group (n=16) ascertained for the (Pennington, Groisser and Welsh) 1993 study may simply be atypical of most individuals with RD+ADHD for reasons that we are unable to determine" (p. 169).

A routine collection of a very large amount of data on all children being evaluated is the approach that was adopted by the NMI clinic. Together with its advantages, this approach also has disadvantages. One disadvantage, which became evident through the amount of missing data, was that in the assessment battery there were several measures of each of the neuropsychological domains. In practice and with regard to the requirements the clinical assessment, all the domains need not be assessed with such a degree of detail for all the children. At the same time, for some children, some abilities required to be assessed in more detail, with measures that were not intended to be implemented with all children. In the case of the NMI clinic, the pre-planned assessment battery was rather extensive, and at the same time, did not fully address the questions set by the clinical assessment. This highlights the difficulty of reaching a suitable compromise between clinical and research needs.

Routine data collection requires decisions concerning the assessment methods to be used, and such decisions can only be based on the knowledge available at the commencement of data collection. As the clinical data collection spans several years, it is very likely that increasing knowledge in the field outdates some of the measures used, as well as introduces new important measures that were not originally included in the test battery. A good example of the effects of increasing knowledge on the new measures adopted for use at the NMI-clinic were measures of phonological processing abilities. At the time of planning and the foundation of the clinic in the mid 80's, the importance of phonological abilities was not fully acknowledged in the neuropsychological research of learning disabilities, although phonological measures had been used in studies focusing on beginning reading skills and difficulties in learning to read. As the importance of phonological abilities in relation to reading disability became evident, and as Finnish measures were developed, the measures were included in (and some others excluded from) the assessment battery. Naturally, the changes in the measures resulted (in terms of data available for research) in a reduction in the number of usable measures. This problem is likely to be present in all attempts to follow pre-planned data collection over a long period of time.

There are also other ways of combining clinical assessment and research than those previously applied in the NMI-clinic and reported here. One promising approach is to start working from more precise research questions from the outset of data collection. In this model, the research questions are formulated beforehand, and the children referred to the clinic are assigned to different research designs according to pre-planned criteria. As a consequence, the assessment methods used vary somewhat from child to child, depending on the research design he/she belongs to. In this model, the child is administered with a wide range of neuropsychological assessments with more detailed assessment based on the individual needs, and on the research design to which he/she belongs. Ideally, this procedure can be completed more effectively than administering a large test battery to all children, and in addition to the individually-selected measures.

This kind of approach to combine clinical work and research data collection has been applied by Hynd and his colleagues, resulting in several studies on ADHD (e.g. Hynd, Lorys, Semrud-Clikeman, Nieves, Huettner & Lahey, 1991; Lahey, Schaughency, Hynd, Carlson & Nieves, 1997; Marshall, Hynd, Handwerk & Hall, 1997; Schaughency, Lahey, Hynd, Stone, Piancentini & Frick, 1989). This approach brings clinical evaluation and research even closer together, the relevance of the assessment methods is inspected from both clinical and research angles. The approach can naturally be applied to a variety of research questions, and may well be a worthwhile direction to adopt when combining clinical assessment and research data collection.

Another way of combining clinical and research work is to recruit children referred to the clinic for further studies. An example of this kind of study is the ongoing co-operation project between NMI-clinic and the University of Jyväskylä. The project focuses on the state regulation deficit hypothesis of ADHD (e.g. van der Meere, 1996). In this project, children referred to the NMI-clinic are screened, and if they fulfil the inclusion criteria of the project, they are offered the possibility to participate. As the measures used in the project take several hours to complete, it is not possible to include them in the general clinical assessment for all children, especially as it is likely that the measures would not give clinically relevant information for most of the children. Thus, the NMI-clinic is utilised as a source to obtain participants for research and certainly, such projects could be more widely applied.

Combining clinical work and research has long traditions in the area of the neuropsychology of developmental disorders. A practical application of the combination has been to use data collected from the clinical setting in research. Combining clinical assessment and research data collection is not unproblematic since the motivations for the two are not equal and indeed, are often competitive. However, previous studies, as well as studies reported in this dissertation, have shown that the combination can be successful, and most importantly such a combination can advance the knowledge on the area. When combining these two, the starting point is to be aware of the unavoidable properties of clinical datasets, and of the different aims of clinical assessment

and research data collection. Acknowledging the limitations and the properties of clinical data helps to overcome the assessment problems and leads to an even more fruitful usage of the combination of clinical work and research.

YHTEENVETO

Neuropsykologisen asiakastyön ja tutkimuksen yhdistämisellä on pitkät perinteet oppimisvaikeuksien neuropsykologisen taustan selvittämisessä. Asiakastyö ja tutkimus eroavat kuitenkin toisistaan merkittävillä tavoilla, ja nämä eroavuudet vaikuttavat asiakastyössä kerätyn aineiston käyttömahdollisuuksiin tutkimuksessa. Asiakastyössä kerättyjen aineistojen tutkimuskäyttömahdollisuudet ovat toisaalta rajalliset, toisaalta sellaisilla aineistoilla on joitain erityisiä etuja. Tässä tutkimuksessa tarkasteltiin Niilo Mäki Instituutin lastentutkimusklinikalla (NMI-klinikalla) kootun aineiston avulla asiakastyön yhteydessä kerättyjen aineistojen käyttömahdollisuuksia oppimisvaikeustutkimuksessa.

Neuropsykologiseen arviointiin oppimisvaikeuksien vuoksi ohjautuneet lapset eivät muodosta edustavaa otosta lapsista, joilla on oppimisvaikeuksia. Aineiston valikoituneisuuteen vaikuttavat tekijät tunnetaan huonosti, ja useimmissa tapauksissa niiden selvittäminen on hyvin vaikeaa tai mahdotonta. Asiakastyössä kerätyistä aineistoista saatuja tutkimustuloksia tarkasteltaessa on muistettava aineiston puutteellinen edustavuus, ja tuloksia on yleistettävä varoen. Yksi asiakastyössä kerättyjen aineistojen tunnettu piirre on niiden lasten suuri osuus joilla on useita erilaisia ongelmia. Komorbiditeetin (ongelmien päällekkäisyyden) yleisyys, jos sitä ei huomioida, vaikeuttaa saatujen tutkimustulosten tulkitsemista. Toisaalta ne tarjoavat mahdollisuuden tutkia komorbiditeettiin liittyviä tekijöitä.

Tutkimuksessa tarkasteltiin lukemisvaikeuden (LV) ja tarkkaavaisuushäiriön (TH) välistä komorbiditeettia. Tarkastelun kohteena oli fenokopio-hypoteesi LV:n ja TH:n komorbiditeetista. Fenokopio-hypoteesin mukaan komorbidiin LV-TH:öön ei liity pelkään TH:öön tyypillisesti liittyviä heikkouksia toiminnanohjauksen taidoissa, vaan komorbidien LV-TH-lasten ensisijainen ongelma on LV ja siihen tyypillisesti liittyvät heikkoudet tietyissä kielellisissä taidoissa. Fenokopio-hypoteesin mukaan TH:tä määrittävät käyttäytymispiirteet ovat toissijaisia LV-TH lapsilla ja seurausta koetusta LV:sta. Tutkimuksessa todettiin oikeaksi se oletus, että lapsilla, joilla on pelkästään LV, on heikkouksia nopean nimeämisen taidoissa. Oletettuja heikkouksia toiminnanohjauksen taidoissa ei havaittu lapsilla, joilla oli pelkästään TH. Komorbideilla LV-TH-lapsilla oli yhtä heikot nopean nimeämisen taidot kuin lapsilla, joilla oli pelkästään LV; he suoriutuivat heikosti myös toiminnanohjauksen tehtävissä. Tulokset eivät tue fenokopio-hypoteesia komorbidista LV-TH:stä, mutta osoittavat, että lukemisvaikeuteen liittyy vaikeuksia nopeassa nimeämisessä riippumatta siitä, liittyykö LV:een TH.

Pyrkimyksenä NMI-klinikalla oli yhdistää asiakastyö ja tutkimusaineiston keruu siten, että kaikista klinikalla käyvistä lapsista kerätään samat tausta- ja tutkimustiedot. Asiakastyön luonteeseen kuuluu, että lapsen tarpeet ja lähetteen kysymyksenasettelut ovat tärkeämpiä kuin tutkimusaineiston keruu. Sen vuoksi lähes kaikkien lasten ja arviointimenetelmien havainnoissa oli puutteita. Imputoinnilla tarkoitetaan puuttuvien havaintojen korvaamista jonkin mallin

tuottamilla arvoilla. Puuttuvien havaintojen ongelmaa on vain harvoin käsitelty syvällisesti asiakastyön yhteydessä kerätyissä neuropsykologisissa aineistoissa, ja imputointimenetelmiä ei ole sovellettu vastaaviin aineistoihin. Tutkimuksessa tarkasteltiin neljän erilaisen imputointimenetelmän tarkkuutta puuttuvien havaintojen ennustamisessa ja niiden tuottamia muutoksia muuttujien keskiarvoihin, variansseihin sekä muuttujien välisiin korrelaatioihin. Mainittujen kriteerien perusteella valittiin menetelmä, jolla aineistossa olleet puuttuvat havainnot täydennettiin.

Asiakastyönä tehtävä tutkimus kattaa laaja-alaisesti neuropsykologiset toiminnot, ja tämän vuoksi asiakastyön yhteydessä kerätyt aineistot ovat erityisen hyödyllisiä selvitettäessä neuropsykologisten testien ominaisuuksia ja keskinäisiä yhteyksiä. Nopean sarjallisen nimeämisen (NSN) tehtävissä vaatimuksena on nimetä sarjallisesti esitettyjä tuttuja visuaalisia ärsykkeitä mahdollisimman nopeasti. NSN:n vaikeuksien yhteydet LV:een on luotettavasti osoitettu. Tutkimuksessa tarkasteltiin erilaisten NSN:n tehtävien keskinäisiä yhteyksiä, ja havaittiin, että erilaisia ärsykkeitä sisältävät ja eri tavoin järjestetyt NSN:n tehtävät mittaavat osittain erilaisia taitoja. NSN:stä selittäviä tekijöitä olivat prosessointinopeus, fonologiset taidot, hienomotorinen nopeus ja kielellinen sujuvuus. Tulokset osoittavat, että suoritus NSN:n tehtävissä on riippuvainen useista eri taidoista, ja se antaa aiheen tarkastella tarkemmin näitä taitoja suhteessa esimerkiksi LV:een.

Trail Making -testin B-osassa (TMT-B) tehtävänä on seurata mahdollisimman nopeasti yhtä aikaa numero- ja aakkossarjoja ja vaihtaa joustavasti niiden välillä. A-osassa (TMT-A) pitää seurata vain numerosarjaa, ja sitä pidetään TMT-B:ssä tarvittavien taustataitojen mittarina. TMT-B:tä käytetään yleisesti toiminnanohjauksen taitojen mittarina. Tutkimuksessa tarkasteltiin lasten, joilla on LV, suoriutumista TMT-A:ssa ja B:ssä suhteessa verrokkilasten suoriutumiseen. Lapset, joilla oli LV, suoriutuivat verrokkeja hitaammin TMT-B:ssä, mutta eivät TMT-A:ssa. Ryhmien välinen ero säilyi, kun TMT-B:n suoritusta verrattiin TMT-A:n suoritukseen. Tutkimusta varten oli luotu tehtävä, joka oli muuten TMT-A:n kaltainen mutta se mittasi aakkossarjan seuraamista (TMT-A-aak). Lapset, joilla oli LV, olivat siinä verrokkilapsia hitaampia. Edelleen, ryhmien välinen ero TMT-B:ssä selittyi TMT-A-aak-tehtävässä suoriutumisessa olleilla eroilla. Tulokset osoittavat, että nopeus, jolla lapsi kykenee seuraamaan aakkossarjaa, on olennainen kontrolloitava tekijä, kun TMT-B:n avulla arvioidaan lapsia, joilla mahdollisesti on LV.

Asiakastyön ja tutkimuksen yhdistäminen vaatii niiden välisten erojen analysointia ja tunnistamista. Sekä asiakastyötä että tutkimusta hyödyttävään lopputulokseen päästään hyväksymällä asiakastyön yhteydessä kerättyjen aineistojen rajoitukset ja hyödyntämällä niiden vahvuuksia tutkimuksia suunniteltaessa.

REFERENCES

Achenbach, T.M. & Edelbrock, C. (1986a). *Manual for the Child Behaviour Checklist and Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M. & Edelbrock, C. (1986b). *Manual for the Teacher Report Form and Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M. (1991a). *Manual for the Child Behaviour Checklist/4-18 and 1991 Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Achenbach, T.M. (1991b). *Manual for the Teacher Report Form and 1991 Profile.* Burlington, VT: University of Vermont, Department of Psychiatry.

Ahonen, T. Leppäsaari, T., Lamminmäki, T., Poikkeus, A.-M., & Siiskonen, T. (1997). Social interactional factors as moderator variables in developmental language disorders: A case study of monozygotic twins. *Logopedics, Phoniatrics, & Vocology, 22*, 79-86.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Author.

Angold, A. Costello, E.J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*, 57-87.

Benton, A.L. (1959). *Right-left discrimination and finger location: development and pathology*. New York: Hoeber-Harper.

Benton, A. L., Hamsher, K. deS., & Shian, A. B. (1994). *Multilingual Aphasia Examination III: Manual of instructions* (3rd ed.). Iowa City, IA: AJA Associates, Inc.

Benton, A.L., Hamsher, K. deS., Varney, N.R. & Spreen, O. (1983). *Contributions to neuropsychological assessment*. New York: Oxford University Press.

Benton, A.L., Hannay, H.J. & Varney, N.R. (1975): Visual perception of line direction in patients with unilateral cerebral disease. *Neuropsychologia, 12*, 183-191.

Beery, K. E. (1989). *Developmental test of visuo-motor integration. Administration, scoring and teaching manual* (3rd edition). Cleveland, OH: Modern Curriculum Press.

Boll, T.J. & Reitan, R.M. (1972). Motor and tactile-perceptual deficits in brain-damaged children. *Perceptual and Motor Skills, 34*, 343-350.

Brown, S.J., Rourke, B.P., & Cicchetti, D.V. (1989). Reliability of tests and measures used in the neuropsychological assessment of children. *The Clinical Neuropsychologist, 3*, 4, 353-368.

Buschke, H., & Fuld, P. A. (1974). Evaluating storage, retention, and retrieval in disordered memory and learning. *Neurology, 11*, 1019-1025.

Caron, C. & Rutter, M. (1991). Comorbidity in child psychopathology; Concepts, issues, and research strategies. *Journal of Child Psychology and Psychiatry, 32*, 1063-1080.

DeFries, J.C. (1989). Gender ratios in reading-disabled children and their affected relatives: A commentary. *Journal of Learning Disabilities*, 22, 544-555.

Denckla, M.B. (1994). Measurement of executive function. In Lyon, G.R. (Ed.), *Frames of reference for the assessment of learning disabilities* (pp. 117-142). Baltimore, ML: Paul H. Brookes.

Denckla, M.B. (1996). A theory and model of executive function: A neuropsychological perspective. In G.R. Lyon & N.A. Krasnegor (Eds.), *Attention, memory and executive function* (pp. 263-278). Baltimore, ML: Paul H. Brookes.

Denckla, M.B., & Rudel, R.G. (1976). Rapid 'automatized' naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia, 14*, 471-479.

DeRenzi, E., & Faglioni, P. (1978). Normative data and screening power of a shortened version of the Token Test. *Cortex, 14*, 14-49.

Dunn, L.M., & Dunn, L.M. (1981). *Peabody Picture Vocabulary Test – Revised*. Circle Pines, MN: American Guidance Service.

Dykman, R., & Ackerman, P.T. (1991). ADD and specific reading disability: Separate but often overlapping disorders. *Journal of Learning Disabilities, 24*, 96-103.

Epstein, M. A., Shaywitz, S.E., Shaywitz, B.A., & Woolston, J.L. (1991). The boundaries of attention deficit disorder. *Journal of Learning Disabilities, 24*, 78-86.

Fisk. J.L. & Rourke, B.P. (1987). WISC/WISC-R comparisons in a learning-disabled population: Equivalence of summary IQ measures. *The Clinical Neuropsychologist, 1*, 47-50.

Francis, D.J., Fletcher, J.M., & Rourke, B.P. (1988). Discriminant validity of lateral sensorimotor tests in children. *Journal of Clinical and Experimental Neuropsychology, 10*, 779-799.

Fuerst, D.R., Fisk, J.L., & Rourke, B.P. (1989). Psychosocial functioning of learning-disabled children: Replicability of statistically derived subtypes. *Journal of Consulting and Clinical Psychology, 57*, 275-280.

Fuerst, D.R. & Rourke, B.P. (1993). Psychosocial functioning of children: Relations between personality subtypes and academic achievement. *Journal of Abnormal Child Psychology, 21*, 597-607.

Gaub, M., & Carlson, C.L. (1997). Gender differences in ADHD: A meta analysis and critical review. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 1036-1045.

Halperin, J.M., Newcorn, J.H., Koda, V.H., Pick, L., McKay, K.E., & Knott, P. (1997). Noradrenergic mechanisms in ADHD children with and without reading disabilities: A replication and extension. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*, 1688-1697.

Heaton, R. K. (1981). *Wisconsin Card Sorting Test Manual*. Odessa, FL: Psychological Assessment Resources, Inc.

Hynd, G.W., Lorys, A.R., Semrud-Clikeman, M., Nieves, N., Huettner, M.I.S., & Lahey, B.B. (1991). Attention deficit disorder without hyperactivity: A distinct behavioural and neurocognitive syndrome. *Journal of Child Neurology, 6 (Supplement)*, S36-S43.

Jöreskog, K., Sörbom, D., du Toit, S., & du Toit, M. (1999). *LISREL 8: New statistical Features*. Chicago, IL: Scientific Software International.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test (Revised 60-item version)*. Philadelphia, PA: Lea & Febiger.

Kaufman, A. S. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.

Kimura, D. (1963). Right temporal lobe damage. *Archives of Neurology, 8*, 264-271.

Kirk, S.A., McCarthy, J.J., & Kirk, W.D. (1968). *Illinois test of Psycholinguistic Abilities. Examiner's manual*. Urbana, IL: University of Illinois Press.

Knights, R.M. & Moule, A.D. (1968). Normative data on Motor Steadiness Battery for children. *Perceptual and Motor Skills, 26*, 643-650.

Knights & Norwood (1979). *A Neuropsychological test battery for children: Examiner's manual*. Ottawa, Canada: Psychological Consultants Inc.

Korkman, M. (1980). NEPS – Lasten neuropsykologinen tutkimus. Helsinki, Finland: Psykologien Kustannus Oy.

Korkman, M., Kirk, U., & Kemp, S. (1998). *NEPSY. A Developmental Neuropsychological Assessment. Manual*. San Antonio, TX: The Psychological Corporation.

Kuusinen, J., Blåfield, L. (1972). *ITPA:n teoria, ominaisuudet ja käyttö*. Kasvatustiedeiden tutkimuslaitoksen julkaisuja 156. Jyväskylä, Finland: Kasvatustiedeiden tutkimuslaitos.

Lahey, B.B., Schaughency, E.A., Hynd, G.W., Carlson, C.L., & Nieves, N. (1987). Attention deficit disorder with and without hyperactivity: Comparison of behavioural characteristics of clinic-referred children. *Journal of the American Academy of Child and Adolescent Psychiatry, 26*, 718-723.

Lamminmäki, T., Leppäsaari, T., Siiskonen, T., Ahonen T., & Lyytinen H. (1994). XYY syndrome and cognitive profile - A diagnostic single case. *Scandinavian Journal of Logopedics & Phoniatrics, 19*, 61-67.

Light, J.G., Pennington, B.F., Gilger, J.W., & DeFries, J.C. (1995). Reading disability and hyperactivity disorder: Evidence for a common genetic etiology. *Developmental Neuropsychology, 11*, 323-335.

Luria, A.R. (1980). *Higher Cortical Functions in Man (2nd edition)*. New York: Basic Books.

Lyytinen, P. (1987). Cognitive skills and Finnish language inflection. *Scandinavian Journal of Psychology, 28*, 304-312.

Marshall, R.M., Hynd, G.W., Handwerk, M.J., & Hall, J. (1997). Academic underachievement in ADHD subtypes. *Journal of Learning Disabilities, 30*, 635-642.

van der Meere, J.J. (1996). The role of attention. In S. Sandberg, (Ed.), *Hyperactivity Disorders in Childhood. Cambridge Monographs in Child and*

Adolescent Psychiatry (pp. 60-109). Cambrigde, UK: Cambridge University Press.

Mäntynen, H., Poikkeus, A.-M., Ahonen, T., Aro, T., & Korkman, M. (submitted). Clinical significance of test item refusals among young children. *Child Neuropsychology*.

Nici, J. & Reitan, R.M. (1987). Patterns of neuropsychological ability in brain-damaged versus normal children. *Journal of Clinical and Consulting Psychology, 54*, 542-545.

Nigg, J.T., Hinshaw, S.P., Carte, S.E., & Treuting, J. (1998). Neuropsychological correlates of childhood attention-deficit/hyperactivity disorder: Explainable by comorbid disruptive behaviour or reading problems? *Journal of Abnormal Psychology, 107*, 468-480.

Niilo Mäki Institute. (1992). *Neuropsychological and achievement tests: Local normative data for Niilo Mäki Institute-Test Battery*. Jyväskylä, Finland: Author.

Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia, 9*, 97-113.

Pennington, B.F., Groisser, D., & Welsh, M.C. (1993). Contrasting cognitive deficits in attention deficit hyperactivity disorder versus reading disability. *Developmental Psychology, 29*, 511-523.

Raven, J. C. (1965). *Guide to using the Coloured Progressive Matrices*. London: H.K. Lewis.

Reader, M.J., Harris, E.L., Schuerholz, L.J., & Denckla, M.B. (1994). Attention deficit hyperactivity disorder and executive dysfunction. *Developmental Neuropsychology, 10*, 493-512.

Reed, H.B.C., Reitan, R.M., & Klove, H. (1965). The influence of cerebral lesions on psychological test performances of older children. *Journal of Consulting Psychology, 29*, 247-251.

Reitan, R.M. (1971). Trail Making Test results for normal and brain-damaged children. *Perceptual and Motor Skills, 33*, 575-581.

Reitan, R.M. & Davidson, L.A. (1974). *Clinical neuropsychology: Current status and applications*. New York: Hemisphere.

Reitan, R.M. & Herring, S. (1985). A short screening device for identification of cerebral dysfunction in children. *Journal of Consulting and Clinical Psychology, 41*, 643-650.

Reitan, R.M., & Wolfson, D. (1985). *The Halstead-Reitan Neuropsychological Test Battery*. Tucson, AZ: Neuropsychology Press.

Reitan, R.M. & Wolfson, D. (1988). Neuropsychological functions of learning-disabled, brain-damaged, and normal children. *The Clinical Neuropsychologist, 3*, 70-78.

Reitan, R.M. & Wolfson, D. (1992). *Neuropsychological Evaluation of Older Children*. South Tucson, AZ: Neuropsychology Press.

Roth, P.L. (1994). Missing data: A conceptual review for applied psychologists. *Personnel Psychology, 47*, 537-560.

Rourke, B.P. (1989). *Nonverbal Learning Disabilities: The Syndrome and the Model.* New York, NY: The Guilford Press.

Rourke, B.P. & Fuerst, D.R. (1992). Psychosocial dimensions of learning disability subtypes: Neuropsychological studies in the Windsor laboratory. *School Psychology Review, 21,* 361-374.

Sarazin, F.F. & Spreen, O. (1986). Fifteen-year stability of some neuropsychological tests in learning disabled subjects with and without neurological impairment. *Journal of Clinical and Experimental Neuropsychology, 8,* 190-200.

Schaughency, E.A., Lahey, B.B., Hynd, G.W., Stone, P.A., Piacentini, J.C., & Frick, P.J. (1989). Neuropsychological test performance and the attention deficit disorders: Clinical utility of the Luria-Nebraska Neuropsychological Battery--Children's Revision. *Journal of Consulting and Clinical Psychology, 57,* 112-116.

Schulte, A.C., Conners, C.K., & Osborne, S.S. (1999). Linkages between attention deficit disorders and reading disability. In D.D. Duane (Ed.), *Reading and Attention Disorders: Neurobiological Correlates* (pp. 161-184). Timonium, ML: York Press.

Selz, M., & Reitan, R.M. (1979). Comparative test performance of normal, learning disabled and brain-damaged older children. *Journal of Nervous and Mental Disease, 167,* 298-302.

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data.* London: Chapman & Hall.

Shaywitz, B.A., Shaywitz, S.E., & Fletcher, J.M (1995). Defining and classifying learning disabilities and attention-deficit/hyperactivity disorder. *Journal of Child Neurology, 10,* S50-S57.

Shaywitz, S.E., Shaywitz, B.A., Fletcher, J.M., & Escobar, M. (1990). Prevalence of reading disability in boys and girls. *Journal of the American Medical Association, 264,* 998-1002.

Smeding, H.M.M. & de Koning, I. (2000). Frontotemporal dementia and neuropsychology: The value of missing values. *Journal of Neurology, Neurosurgery, and Psychiatry, 68,* 726-730.

Snow, J.H. (1998a). Clinical use of the Benton Visual Retention Test for children and adolescents with learning disabilities. *Archives of Clinical Neuropsychology, 13,* 629-636.

Snow, J.H. (1998b). Developmental patterns and use of the Wisconsin Card Sorting Test for Children and adolescents with learning disabilities. *Child Neuropsychology, 4,* 89-97.

Snow, J.H., English, R., & Lange, B. (1992). Clinical utility of the Selective Reminding Test-children's version. *Journal of Psychoeducational Assessment, 10,* 153-160.

Spreen, O. & Haaf, R.G. (1996). Empirically derived learning disability subtypes: A replication attempt and longitudinal patterns over 15 years. *Journal of Learning Disabilities, 19,* 170-180.

Stevenson, J., Pennington, B.F., Gilger, J.W., DeFries, J.C., & Gillis, J.J. (1993). Hyperactivity and spelling disability: Testing for shared genetic aetiology. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 34,* 1137-1152.

Tiffin, J. (1968). *The Purdue Pegboard.* Chicago, IL: Research Associates, Inc.

Wagner, R.K., Torgesen, J.K., Laughon, P., Simmons, K., & Rashotte, C.A. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology, 85,* 83-103.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children (rev. ed.).* New York: Psychological Corporation.

Wecshler, D. (1991). *Wechsler Intelligence Scale for Children-III.* San Antonio, TX: Psychological Corporation.

Willcutt, E.G., Pennington, B.F., Boada, R., Ogline, J.S., Tunick, R.A., Chhabildas, N.A., & Olson, R.K. (2001). A comparison of the cognitive deficits in reading disability and attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology, 110,* 157-172.

Wolf, M. (1986). Rapid alternating stimulus naming in the developmental dyslexias. *Brain and Language, 27,* 360-379.

Wolf, M., & Bowers, P.G., & Biddle, K. (2000). Naming-speed processes, timing, and reading: A conceptual review. *Journal of Learning Disabilities, 33,* 387-407.

Woodward, L., Dowdney, L., & Taylor, E. (1997). Child and family factors influencing the clinical referral of children with hyperactivity: A research note. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 38,* 479-485.